

DOCUMENT RESUME

ED 249 471

CS 007 800

AUTHOR Weinshank, Annette; And Others
TITLE Learning from Experience to Improve Outcomes in Reading: A Case Study. Research Series No. 149.
INSTITUTION Michigan State Univ., East Lansing. Inst. for Research on Teaching.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
REPORT NO IRT-RS-149
PUB DATE Jun 84
CONTRACT 400-81-0014
NOTE 17p.
AVAILABLE FROM Institute for Research on Teaching, College of Education, Michigan State University, 252 Erickson Hall, East Lansing, MI 48824 (\$2.50).
PUB TYPE Reports - Research/Technical (143)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Diagnostic Teaching; Grade 7; Instructional Improvement; Junior High Schools; Listening Comprehension; Oral Reading; Reading Achievement; Reading Comprehension; *Reading Diagnosis; *Reading Improvement; *Reading Instruction; Reading Research; *Remedial Reading; Silent Reading; Teacher Effectiveness; Word Recognition

ABSTRACT

A study of the requisite processes for establishing diagnostic validity in reading was conducted in a seventh grade remedial reading classroom. The study's objectives were to (1) collect reliable diagnostic data on each student's performance in word recognition, oral reading, silent reading comprehension, and listening comprehension before and after the year's instruction; (2) document the teacher's instructional practices; (3) link outcomes with instruction; and (4) return diagnostic and outcome information to the teacher. On the basis of this information, the teacher made instructional adjustments for the next year's program in the areas of word recognition and oral reading. The diagnostic results for the following year showed improvement in student achievement in those areas. The results also indicated that an unobtrusive system for establishing diagnostic validity can be successfully introduced into the classroom. (Author/HTH)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

X This document has been reproduced as
received from the person or organization
originating it.
☐ Minor changes have been made to improve
reproduction quality.

* Points of view or opinions stated in this docu-
ment do not necessarily represent official NIE
position or policy.

ED249471

Research Series No. 149

LEARNING FROM EXPERIENCE
TO IMPROVE OUTCOMES IN READING:
A CASE STUDY

Annette Weinshank, Ruth Polin,
and Christian Wagner

Published By

The Institute for Research on Teaching
252 Erickson Hall
Michigan State University
East Lansing, Michigan 48824-1034

Printed and Distributed
by the
College of Education
Michigan State University

June 1984

This work is sponsored in part by the Institute for Research on Teaching,
College of Education, Michigan State University. The Institute for Research
on Teaching is funded primarily by the Program for Teaching and Instruction of
the National Institute of Education, United States Department of Education.
The opinions expressed in this publication do not necessarily reflect the
position, policy, or endorsement of the National Institute of Education.
(Contract No. 400-81-0014)

008 L00 50

Institute for Research on Teaching

The Institute for Research on Teaching was founded at Michigan State University in 1976 by the National Institute of Education. Following a nationwide competition in 1981, the NIE awarded a second contract to the IRT, extending work through 1984. Funding is also received from other agencies and foundations for individual research projects.

The IRT conducts major research projects aimed at improving classroom teaching, including studies of classroom management strategies, student socialization, the diagnosis and remediation of reading difficulties, and teacher education. IRT researchers are also examining the teaching of specific school subjects such as reading, writing, general mathematics, and science, and are seeking to understand how factors outside the classroom affect teacher decision making.

Researchers from such diverse disciplines as educational psychology, anthropology, sociology, and philosophy cooperate in conducting IRT research. They join forces with public school teachers, who work at the IRT as half-time collaborators in research, helping to design and plan studies, collect data, analyze and interpret results, and disseminate findings.

The IRT publishes research reports, occasional papers, conference proceedings, a newsletter for practitioners, and lists and catalogs of IRT publications. For more information, to receive a list or catalog, and/or to be placed on the IRT mailing list to receive the newsletter, please write to the IRT Editor, Institute for Research on Teaching, 252 Erickson Hall, Michigan State University, East Lansing, Michigan 48824-1034.

Co-Directors: Jere E. Brophy and Andrew C. Porter

Associate Directors: Judith E. Lanier and Richard S. Prawat

Editorial Staff

Editor: Janet Eaton

Assistant Editor: Patricia Nischan

Abstract

Learning from experience to improve the teaching of reading depends on teachers having access to usable information about the effect of their practice on student outcomes. Reliable diagnostic classification, documentation of instruction, and the measurement of outcomes are needed to establish a system for verifying the links that connect diagnosis with instruction and outcomes. A study of the requisite processes for establishing diagnostic validity in reading was conducted in a seventh-grade remedial-reading classroom. The study's objectives were (1) to collect reliable diagnostic data on each student's performance in word recognition, oral reading, silent reading comprehension, and listening comprehension before and after the year's instruction; (2) to document the teacher's instructional practices; (3) to link outcomes with instruction; and (4) to return diagnostic and outcome information to the teacher. On the basis of this information, the teacher made instructional decisions for the next year's program in the areas of word recognition and oral reading. Outcome results the following year showed improvement in student achievement in those areas. The results indicate that an unobtrusive system for establishing diagnostic validity can be successfully introduced into the classroom.

LEARNING FROM EXPERIENCE TO IMPROVE OUTCOMES
IN READING: A CASE STUDY¹

Annette Weinshank, Ruth Polin, and Christian Wagner²

How can teachers learn more effectively from their own practice?

Experience may indeed be a valuable teacher, but it has a serious limitation: Classroom events happen too quickly for teachers to aggregate reliably the results of instruction in order to profit from them. Isolation and over-reliance on personal experience further restrict the range of information available to a teacher that can be used to evaluate instructional decisions and actions in light of student achievement outcomes.

The absence of specific information linking student outcomes back to reliable diagnosis and instruction has been a major obstacle to teachers of reading who want to learn systematically from their own experience. Improving student outcomes in reading requires that the teacher receive information that reliably classifies student performance into appropriate diagnostic categories, because assurance of diagnostic reliability is central to evaluating one's instructional practices. If the diagnostic categorizations are unreliable, it is impossible to determine whether instruction was inadequate for the student's problem or whether student performance was incorrectly diagnosed and good instruction was wasted on the wrong problem.

Unfortunately, diagnostic unreliability pervades the field of reading. Only recently has it become possible to remedy the situation and to provide

¹This paper was presented at the American Educational Research Association meetings held April 1984 in New Orleans, Louisiana.

²Annette Weinshank is an IRT teacher collaborator and coordinator of the Outcomes in Reading Project. Ruth Polin is a research associate for that project. Christian Wagner, who teaches in the College of Engineering at Oakland University, serves as the project's consultant.

teachers with reliable diagnostic information. A series of studies on the diagnostic reliability of educational practitioners (Vinsonhaler, Weinshank, Wagner, & Polin, 1983; Weinshank, 1982; Weinshank & Vinsonhaler, 1983) showed that reading and learning disability specialists and classroom teachers did not agree with themselves or with one another in their diagnostic judgments about simulated cases of children with reading problems. A second series of studies (Vinsonhaler, Weinshank, Polin, & Wagner, 1983) showed that the reliability of diagnostic decision making could be improved dramatically through the use of a specific type of training program.

Two important products of the training-study series were (1) a reading diagnostic test battery based primarily on performance in instant word recognition, decoded word recognition, oral reading, silent reading comprehension, and listening comprehension and (2) a computer program for generating a reliable diagnosis given the resulting student performance data. Thus it became possible to provide teachers with reliable diagnostic information about their students' reading performance.

This paper presents a case study of a junior high school remedial reading teacher who was provided with reliable diagnostic information about her students' reading performance and was thus able, for the first time, to learn systematically from her own experience and make changes for the better in her instruction.

A Case Study of Diagnostic Validity

The ability to train practitioners to be more reliable and/or have reliable diagnoses provided them became the basis for a series of outcome studies in the senior author's seventh-grade pull-out remedial reading program. (The senior author is a teacher collaborator with the Institute for Research on

Teaching, conducting research half-time and teaching half-time in the public schools.) The studies were designed to test the methods needed to move from diagnostic reliability to diagnostic validity. Given reliable diagnostic information, what instructional plans do teachers carry out and with what effect on students?

During the study's first year (Wagner, 1982), the diagnostic battery was administered to each student in the remedial-reading program as a pretest and a posttest. Using diagnostic decision aids developed in the training series, the teacher diagnosed each student. Using the same data, the computer program also generated a diagnosis for each student. Agreement between teacher and computer was better than 90%. The remedial instruction that the teacher implemented with her students was documented by her daily notes and interview and observational data collected by project staff. At year's end, gain scores across the diagnostic categories were computed for all students.

The information on student gains confirmed some of the teacher's subjective evaluations and disconfirmed others. Much lower than anticipated gains in oral reading led her to alter instructional procedures for the next year's students. With the help of an instructional aide, the teacher instituted two major changes in her practice. First, she decided to abandon drill on common word families; her expectation that the students would generalize the patterns to unfamiliar words encountered in text was not supported by the performance scores. Instead, she would draw vocabulary words exclusively from the materials the students read. Students would have to master these words to speed and accuracy criteria before reading text. Syllabication strategies and sound-symbol association drill would be embedded within the vocabulary words. The teacher also decided that instead of conducting group oral reading for a

portion of each session, she would require daily oral reading on an individual basis.

Instructional Goals

The teacher said her general instructional goals were improved word recognition and analysis skills within the context of extended, supervised oral reading. She had two major goals: (1) automate analysis and recognition skills in a framework of contextual reading and (2) give the students confidence in their ability to control the flow of print visually and orally. She said that direct instruction in comprehension ran second to these two instructional goals. In her judgment, instructional time constraints coupled with serious deficits in word analysis and fluency dictated a primary emphasis on bringing students up to grade placement on those skills. Comprehension probes were used informally within the oral reading context.

The Reading Program

Five groups of four to six students came to the reading room from their respective classrooms for a 25-minute session four days a week between October and May, for a total of 40 instructional hours. The students moved through a fixed sequence of materials, entering at a level determined by the results of their performance on the word recognition and oral reading subtests of the battery and generally moving through the sequence at their own pace. Sometimes, students were allowed to skip certain steps. The program functioned more like a clinic than a classroom. Students read orally twice during each session. Lists of vocabulary words preceded various subsets of text, and those words had to be mastered to speed and accuracy criteria before the student was allowed to begin oral reading.

The teacher described her program as staged. In Stage 1, students used a linguistically based reading program. This type of program typically starts by presenting single-syllable words made up of highly regular sound-symbol correspondences and then proceeds slowly to more irregular ones. Students were placed in reading texts appropriate to their levels of performance. No worksheets or isolated phonics drills were used. In Stage 2, a small number of syllabication heuristics were presented so that the previously mastered sound-symbol correspondences could be transferred to two- and three-syllable words. Content-based texts were used during this phase. Stage 3 was designed to move students toward fluency and automaticity with texts of increasing difficulty. A variety of basal texts at grade levels 6 and 7 were used for contextual practice.

The teacher recorded oral reading performance throughout the year. Every time a student read, the date, page(s) read, elapsed time, and words miscalled were written on a sheet devoted to the particular text the student was working with at that time. Text read at a very slow rate and/or with excessive miscalls was repeated and documented on a record sheet so that students could note changes in their own performance. After completing an oral reading session, the student continued reading silently alone until called back 5-10 minutes later to read orally again.

The procedures for the second year's study were to (1) collect, using the diagnostic battery, reliable diagnostic data on each student before the year's instruction began and return the performance information to the teacher; (2) repeat diagnostic testing at the end of the year and return the performance information to the teacher; (3) determine student outcomes and return that information to the teacher; and (4) examine the links between instruction and outcomes.

Collecting Reliable Diagnostic Data

At the beginning and end of the year, the diagnostic battery was administered individually by trained testers to the 23 students in the reading program, all of whom began the program reading two or more years below grade placement. The battery included measures of known reliability and provided data on seven variables of interest: (1) word recognition and (2) percent of words correctly identified by decoding (Slosson Oral Reading Test), (3) oral reading of paragraphs based on number of miscalls and (4) mean words per minute (Gray Oral Reading Paragraphs), (5) silent reading comprehension based on percent of passage recalled and (6) mean words per minute (silent reading test of the Durrell Analysis of Reading Difficulty), and (7) listening comprehension based on percent of passage recalled (listening comprehension test of the Durrell Analysis of Reading Difficulty).

Returning Pretest Information to the Teacher

The teacher received pretest diagnostic information in two different forms: (1) numerically, as individual and mean group scores for each of the seven diagnostic categories, and (2) graphically, as performance charts. The four charts (word recognition and percent decoding, oral reading grade equivalent vs. rate, silent reading grade equivalent vs. rate, and listening comprehension grade equivalent) allowed the teacher to see group configurations before and after instruction and to track changes for individual students. Table 1 shows a pretest performance chart for oral reading.

In this class, 87% of the students entered reading orally at three or more years below grade placement. By the end of the year, that figure had dropped to 57%. There was also a trend toward increased reading rate. On the individual level, Student 1014, for example, entered the program reading at a third-grade level and exited reading at an eighth-grade level.

Table 1
Oral Reading: Pretest Performance

		RATE: WORDS PER MINUTE							
		0-60	61-70	71-80	81-90	91-100	101-110	111-120	121-130
GRADE EQUIVALENT	6th					1030			
	5th			1003 1021 1027					
	4th			1012 1032	1024 1005	1009 1033 1017		1028	1034
	3rd	1023	1015 1011	1002 1004 1014	1029	1026	1016		
	2nd			1020					

4-digit number = student I.D.

Posttest and Outcome Information

The diagnostic battery was again administered to all students after the year's instruction was over. Values for each student (posttest score minus pretest score) for each of the diagnostic categories were calculated. Posttest and outcome information was returned to the teacher in numeric and graphic form. The posttest and outcome charts for oral reading are shown in Tables 2 and 3.

Outcome results show the differential effects of the teacher's instruction, at least as measured by the administered diagnostic tests. The students showed strong gains in oral reading accuracy with some signs of increased rate. Their word recognition skills improved, and their use of

Table 2
Oral Reading: Posttest Performance

		RATE: WORDS PER MINUTE							
		0-60	61-70	71-80	81-90	91-100	101-110	111-120	121-130
GRADE EQUIVALENT	10th							1030	
	8th			1014					
	6th			1016 1015	1021	1005 1003 1024	1027		
	5th		1002	1032 1029	1026	1034 1033 1009 1028 1017			
	4th		1004	1012	1020	1023			
	3rd			1011					

4-digit number = student I.D.

Table 3
Oral Reading: Outcomes

		CHANGE IN RATE: WORDS PER MINUTE						
		-30	-20	-10	0	+10	+20	+30
LOSSES/GAINS (YEARS)	+5				1014			
	+4					1030		
	+3		1016			1015		
	+2			1002 1026	1029	1024 1020 1005		
	+1	1034 1028			1004 1009 1032 1017	1033	1003 1021	1027 1023
	0				1012	1011		
	-1							

T gain = 1.7 years

decoding as a strategy for word attack increased. Performance on silent reading comprehension paragraphs was not as encouraging. A trend toward increased silent reading rate suggests that the students may well have been reading both more fluently and more accurately given their oral reading and word recognition performance, but these subskills did not appear to affect comprehension of the test paragraphs. Listening comprehension showed the least change. In general, the students seemed to find the subject matter of the paragraphs as obscure at the end of the year as at the beginning.

To summarize, the outcome results were as follows:

1. Word recognition. The mean gain was 1.3 years in grade level (range = .3-2.3); mean percent of words decoded rose from 6.5% on the pretest to 16% on the posttest.
2. Oral reading. The mean gain, as measured by number of miscalls, was two grade levels; mean words per minute remained essentially unchanged (84 words per minute on the pretest and 87 words per minute on the posttest).
3. Silent reading comprehension. The mean gain was half a grade level (range = 0-3.0); mean words per minute increased 7.4% (100 words per minute on the pretest; 108 on the posttest).
4. Listening comprehension. There was essentially no change in mean performance (grade level 5.0 on the pretest; grade level 5.1 on the posttest).

Linking Instructional Practices to Outcomes

The overall outcome results confirmed the relative efficacy of the teacher's instruction in word recognition and oral reading and its relative lack of impact on silent reading comprehension and listening comprehension. A more fine-grained examination of the links that connect diagnosis with instruction and outcomes is made possible by observing the effects of the instructional program on students in any given pretest cell of a performance chart. If each cell does, in fact, represent a diagnostic category that correctly separates

like from unlike students, then it should be possible to pinpoint differential effects of instruction across categories.

Outcome results for both word recognition and oral reading for students with identical pretest scores were analyzed. Results show that the teacher was most effective with the most disabled of the students in both word recognition and oral reading. In word recognition, the strongest gains were made by the students who entered three or more years below grade placement. In oral reading, the strongest gains were made by students who entered four years below grade placement. Of the 11 students in this category, nine made gains ranging from one to four years. The effects of instruction on oral reading rate are ambiguous. Strong gains in grade-level placement were as often associated with increases in rate as with decreases.

Analyzing word recognition performance on the dimension of percent decoding showed that all students in the 5% decoding category increased their level of decoding irrespective of grade-level placement. However, of the students in the 15% category, only those who entered three or more years below grade placement increased the number of words arrived at via decoding. This appears to further confirm the efficacy of the program for the weakest readers.

An analysis of the teacher's instructional documentation for oral reading provided information, by student, on oral-reading rate, number of words read orally, number of times (trials) a selection was reread, a list of all words that were miscalled, and attendance. Findings for this class include the following:

1. Mean rate for oral reading during instruction was 63 ± 7 words per minute. This was slower than the rate for oral reading paragraphs both on the pretest (84 ± 17) and the posttest (87 ± 12).

2. Total mean number of words read orally during the school year was approximately 26,000. Students who made the strongest gains were below the mean; they read approximately 20,000 words.
3. Mean number of trials was 1.5. Repeated trials (more than 2) were effective in reducing miscalls but ineffective or counter-productive in increasing rate. The teacher has learned that having students read a selection more than twice is an exercise in diminishing returns; students would be better off reading another selection of comparable difficulty.
4. Across all 23 students, 2,478 words were miscalled during oral reading. Of the list words mastered prior to oral reading, only 15% were subsequently missed, suggesting that list mastery probably contributed to fluency.
5. Mean attendance was 82%; most of the students came to most of the sessions.

For the first time in her teaching career, this teacher received specific information that enabled her to link day to day instructional activities with student achievement in reading. On the basis of this information she made several decisions about the next year's program: (1) continue using text-based vocabulary for decoding and syllabication practice, (2) limit repeated reading of a passage to two trials only, (3) eliminate use of the linguistically-based materials except for the very weakest students (stronger students will enter the sequence at Stage 2), and (4) maintain comprehension instruction at the same informal level.

Establishing Diagnostic Validity Within and Across Classrooms

The results indicate that it is possible to introduce into a classroom setting a system for establishing diagnostic validity, that is, establishing the links that connect diagnosis with instruction and student achievement. The procedures are non-intrusive and sensitive to the teacher's regular instructional practices. Receiving information about the results of instruction may or may not persuade a teacher to alter practice, but it does provide a rational basis for decision making. This teacher found that overall, her

instructional program served her goals of improving word recognition and oral reading skills. She learned that her program is more effective for the lowest achieving entering seventh graders than it is for those somewhat less disabled. She has decided to alter her instructional strategy for the stronger entering students but has not decided how (or whether) to expand her goals to include more direct instruction in comprehension.

The teacher can continue to learn from her own teaching practice by repeating the documentation of the diagnostic-treatment-outcome cycle year after year. But the process is slow and there is no comparative information available to her from other teachers who deal with similar students. How do students with similar entering characteristics fare in classroom settings that are similar to or different from hers? How do they fare with similar or different materials and strategies?

Diagnostic validity for students who range across the whole spectrum of reading performance can be established over time only by practitioners willing to learn from their own practice and from that of their colleagues. Learning from experience to improve practice in reading depends on teachers having access to reliable information about the outcomes of their own and their colleagues' instructional practices. In a study currently under way, the authors are extending the procedures described for establishing diagnostic validity to 10 fifth-grade classrooms. A major goal of the research is to document a full range of diagnostic classifications, associated instruction, and student achievement in reading. Such documentation would provide a benchmark against which teachers and researchers could measure the consequences of instruction both within and across classrooms. The practitioners involved in the study will be in a position to learn from and share their own and their

colleagues' experience while simultaneously contributing to the growth of a descriptive, and ultimately prescriptive, data base that connects diagnosis with instruction and outcomes.

References

- Vinsonhaler, J., Weinshank, A., Polin, R., & Wagner, C. (1983). Improving diagnostic reliability in reading through training (Research Series No. 126). East Lansing, MI: Institute for Research on Teaching, Michigan State University.
- Vinsonhaler, J., Weinshank, A., Wagner, C., & Polin, R. (1983, Winter). Diagnosing children with educational problems: Characteristics of reading and learning disabilities specialists and classroom teachers. Reading Research Quarterly, 18(2), 134-164.
- Wagner, C. (1982). Learning from experience: Applying the clinical and epidemiological research paradigms to the study of diagnosis and treatment in reading. Unpublished doctoral dissertation, Michigan State University, East Lansing, MI.
- Weinshank, A. (1982). The reliability of diagnostic and remedial decisions of reading specialists. Journal of Reading Behavior, 14(1), 33-50.
- Weinshank, A., & Vinsonhaler, J. (1983, January). On diagnostic reliability in reading: What's wrong and what can be done? Topics in Learning and Learning Disabilities, 2(4), 43-51.